

①

ETIC FILE COPY

AD-A229 348

STATEMENT OF
WORK
ON THE
GOES I-M
PROJECT

U
E
CO

Destriping GOES Images by Matching Empirical Distribution Functions

M. P. Weinreb, R. Xie,* J. H. Lienesch, and D. S. Crosby†

National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, Washington, DC

The current and future geostationary operational environmental satellites (GOES) of the National Oceanic and Atmospheric Administration (NOAA) are designed to produce visible images of the earth with linear arrays of eight detectors. Because the imaging instruments are not calibrated radiometrically in orbit, differences among instrument gains associated with the different detectors may cause artificial stripes to appear in the images. In the data processing on the ground, the images are "normalized" to remove the stripes. Images from future geostationary satellites, GOES I-M, will be normalized by the method of matching empirical distribution functions (EDFs). In this paper we report on a study of EDF matching with data from GOES-7. The technique was used to generate a normalization look-up table from data taken on 18 May 1988, and the table was applied to image data obtained 2 weeks later, on 1 June 1988. This removed the stripes from the image. The technique is expected to be even more effective with data

from GOES I-M because of improvements in instrumentation.

INTRODUCTION

Both the current and future Geostationary Operational Environmental Satellites (GOES) of the National Oceanic and Atmospheric Administration (NOAA) are designed to carry instruments that image the full disk of the earth, or sections of it, in the visible region of the electromagnetic spectrum. The current satellite, GOES-7 (Ensor, 1978), spins at 100 rpm on a north-south axis. Visible radiation is detected with an array of eight photomultipliers in a north-south line. The field of view from each photomultiplier is approximately 0.8 km sq on the earth at the nadir. With each rotation of the satellite, eight parallel adjacent lines are swept out in the west-to-east direction. After each rotation a scan mirror is stepped to displace the fields of view in the north-south direction. In the early 1990s, the GOES I-M system (Komajda and McKenzie, 1987) is expected to become operational. Instead of spinning, those satellites will be three-axis stabilized and will always present the same side to the earth. Each GOES I-M imager will utilize a linear array of eight detectors, which will be silicon photodiodes instead of photomultipliers, and it will scan both west-to-east and east-to-west. Despite

*Permanent affiliation: Satellite Meteorological Center, Beijing, People's Republic of China.

†Permanent affiliation: The American University, Washington, DC.

Address correspondence to Dr. M. P. Weinreb, NOAA/NESDIS, Code E/RA14—Phys. Br., World Weather Bldg., Rm. 810, Washington, DC 20233.

Received 3 January 1989; revised 19 May 1989.

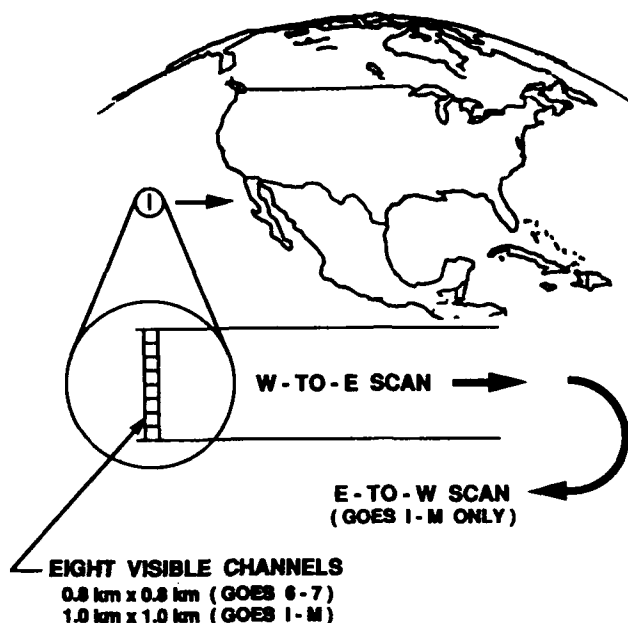


Figure 1. Scanning geometry of GOES imagers (not to scale).

such differences in instrumentation, the current and future GOES imagers will have similar scanning geometries, depicted in Fig. 1.

Henceforth, we will refer to each detector and its associated optics and signal-processing electronics as a channel. Each imaging instrument is then said to have eight visible channels. When stimulated with incident radiation, each channel responds with an output, measured in digital counts. We use the term "gain" to refer to the ratio of the change in the count output to the change in intensity of the incident radiation. The visible channels of the GOES I-M imagers are designed to be linear, i.e., the output will be a linear function of incident intensity, and the gain will be independent of intensity. However, the visible channels of GOES-7 were designed with the output approximately equal to the square root of the incident intensity to take advantage of the increased sensitivity of the photomultipliers at low intensity levels (Ensor, 1978; Bristol, 1975). The photomultipliers, which typically operate in a slightly nonlinear fashion, may also add to the nonlinearity.

Once a GOES is in orbit, its visible channels cannot be calibrated radiometrically, because there is no calibrated source of visible radiation on board. Since the gains (and offsets) of the eight channels may not all be equal and may change with time, artificial east-west stripes can appear in the images. Figure 2 shows striping in an image from

GOES-7. Striping may also affect image data from GOES I-M, although it is expected to be less severe. One reason for this is that the gains of the photodiode detectors of GOES I-M are expected to be more nearly uniform and more stable in time than the gains of the photomultipliers of GOES-7. Also, the data from GOES I-M will be digitized more finely. The 6-bit word structure of GOES-7 will be increased to 10 bits on GOES I-M. The more bits per word, the less significant the striping from quantization error.

In operational data processing at NOAA's National Environmental Satellite, Data, and Information Service (NESDIS), we compensate for channel-to-channel differences in gain with a procedure called normalization, which is applied to the data with look-up tables. It is done automatically, continuously, and in real time at the ground station at Wallops, VA, as the data are received from the satellites and then retransmitted to users. Look-up tables are generated off line (and not in real time) in the main-frame computers at Suitland, MD. For each satellite instrument, a single table is applied to all image data from the entire disk of the earth. Because characteristics of the instruments change with time, new look-up tables are generated occasionally, usually once a week or less often. The current method of generating the operational look-up tables (J. Lienesch, NOAA/NESDIS, personal communication, 1980) has been in use since the mid 1970s. However, NESDIS plans to apply a different method, namely, the matching of empirical distribution functions (EDFs), to generate the tables for GOES I-M, because it is more efficient.

Matching of EDFs is a standard statistical technique and has already been applied to satellite data from the Landsat Multispectral Scanner (MSS) [see, e.g., Horn and Woodham (1979)]. In that work, a normalization look-up table is derived from and applied to the same sector of image data. Our application is different, however, because we are required to derive a table from a single "dependent" sector on a particular day and apply it for weeks afterwards to independent image data from all over the earth's disk.

In preparation for the launch of GOES I, we have been studying the feasibility of the EDF technique by applying it to data from GOES-7, which was the only multichannel satellite imager whose data were conveniently available to us. Data

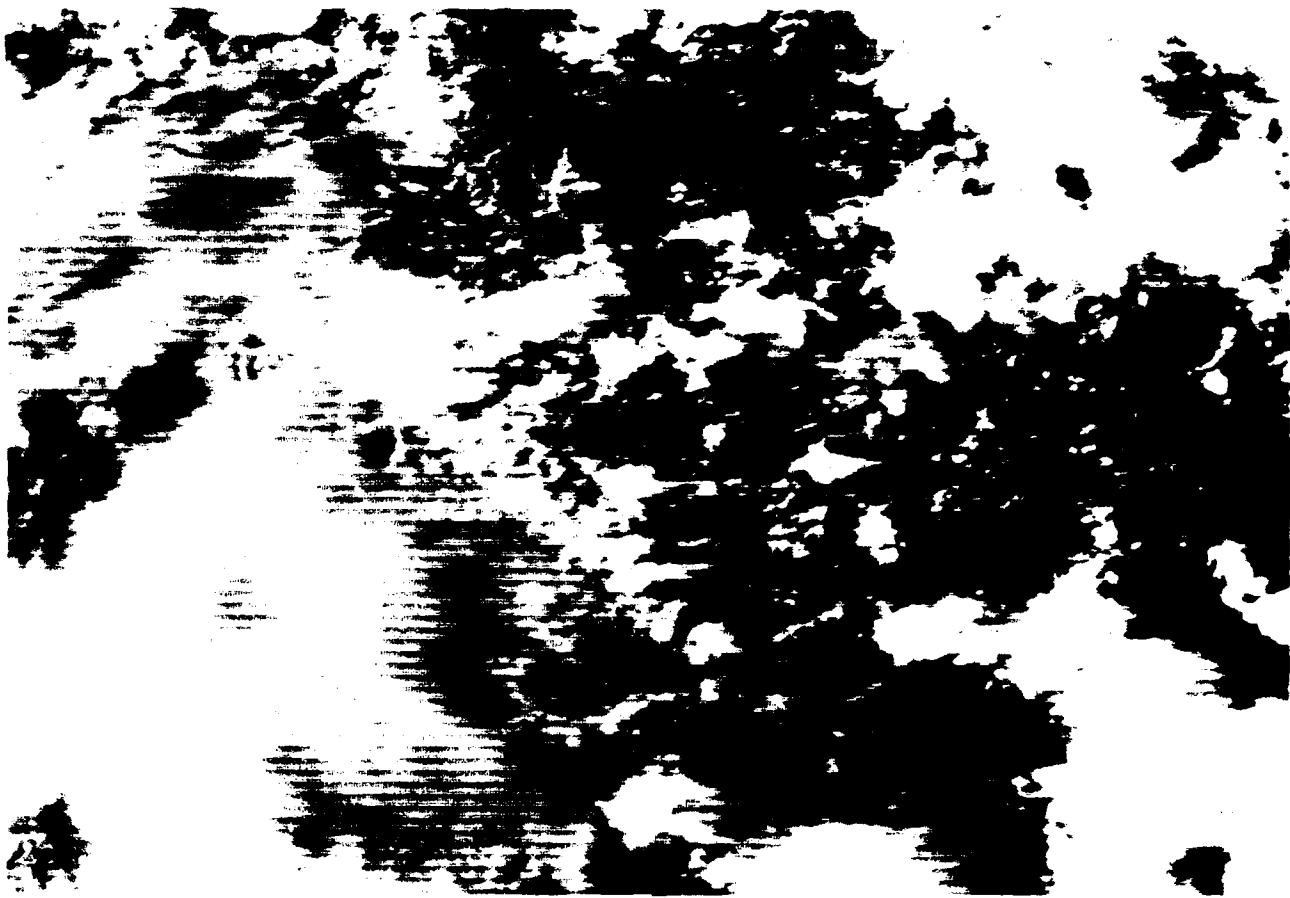


Figure 2. Unnormalized GOES-7 image, 18 May 1988.

from GOES-7 offer a severe test of the method, because it is expected that the GOES I-M sensors will be better behaved and more stable in time than those of GOES-7, as we explained previously. This paper summarizes some of the results from the GOES-7 study. More details can be found in Weinreb et al. (1989).

THEORY

The basic premise is that if several channels view the same scene, their outputs should be equal, and this should be the case regardless of how bright the scene is. In actual application, no two channels ever view the same scene. Instead, we assume that with a large ensemble of measurements, the distribution of the intensity of the earth radiation incident on each detector will be similar (Horn and Woodham, 1979). (In practice, the distributions will not be identical, but the larger the ensemble, the more similar they will become.) With that

assumption, the basic premise becomes that the distributions of the outputs of each channel should be identical.

In our approach, we designate one of the channels as a reference channel. Then the outputs of the other channels are adjusted with the normalization tables so that their distributions are the same as that of the reference channel. The reference channel should be selected on the basis of its relative stability, low noise, and maximal use of the dynamic range of the data system without clipping at either the low or high ends.

To generate a normalization look-up table, we begin by selecting a sample of full-resolution unnormalized earth-scene data covering as much of the range of intensities as possible. For GOES I-M, the area will be rectangular, extending several thousand pixels both east to west and north to south. Corresponding to the incoming radiance from any pixel, the instrument will respond with an output x , in digital counts. One can compile the discrete density function, i.e., the histogram, describing the relative frequency of occurrence of

each possible count value, for each channel. For Channel i , which is the channel to be normalized, let the histogram be $p_i(x)$. An empirical distribution function (EDF) $P_i(x)$ can then be generated; viz.,

$$P_i(x) = \sum_{t=0}^x p_i(t).$$

(The EDF is also known as a cumulative histogram of relative frequency.) The EDF is a nondecreasing function of x , and its maximum value is unity. For convenience, however, we have chosen the maximum value to be 100%; i.e., if the maximum possible output in counts is X , then

$$P_i(X) = 100\%.$$

In these terms, the basic premise says that for each output value x in Channel i , the normalized value x' should satisfy

$$P_r(x') = P_i(x), \quad (1)$$

where the subscript r refers to the reference channel. In practice, not only is P_r nondecreasing, but it is also monotonically increasing as a function of x' in the domain of x' where there are data. Therefore, it can be inverted, yielding the solution for x' ,

$$x' = P_r^{-1}(P_i(x)). \quad (2)$$

When it is applied sequentially for every possible count value x , Eq. (2) generates the normalization look-up table relating each x to an x' . Figure 3 depicts how the procedure is applied in actual practice to generate one entry in the table. The

Figure 3. Illustration of procedure to generate normalization look-up table (see text for explanation).

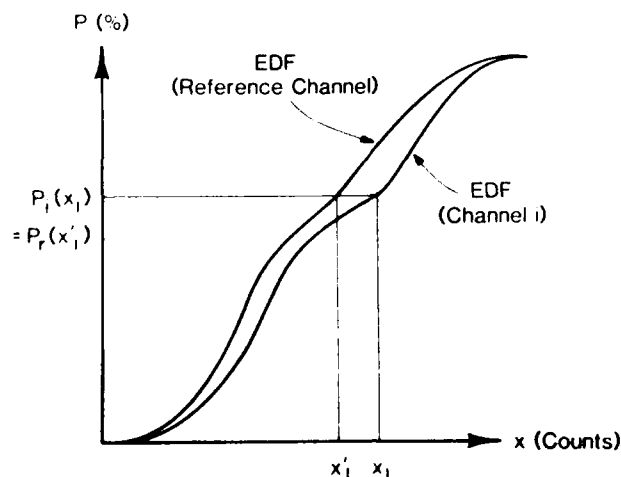
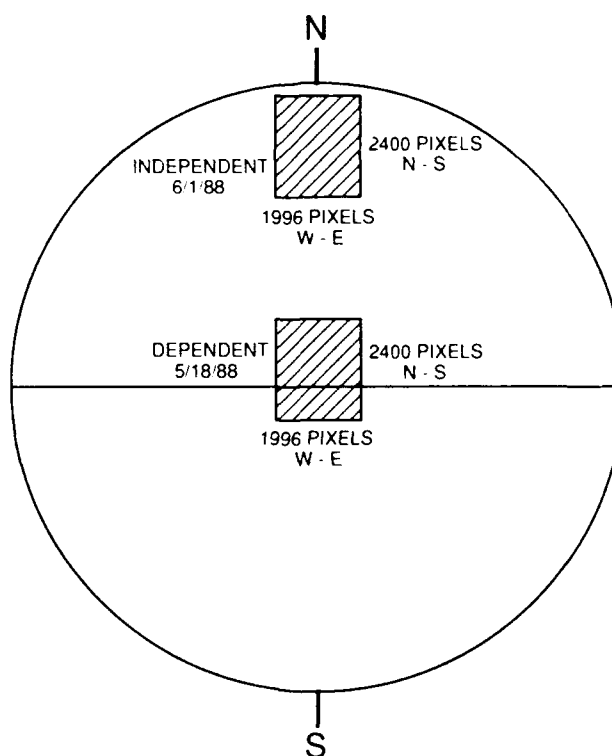


figure depicts idealized EDFs for the reference channel and Channel i . In the figure the EDFs are continuous, but in practice they are discrete, being specified only at integer values of x . To find x'_1 , the normalized count value corresponding to the observed count value of x_1 , the following is the procedure: First, for the count value x_1 in Channel i , find the percentage value from the EDF of Channel i . In the illustration it is $P_i(x_1)$. Then find the point on the reference channel's EDF with the same percentage value. According to Eq. (1), that percentage can also be expressed as $P_r(x'_1)$. Finally, use the EDF of the reference channel to find the normalized count value x'_1 . Since the data are actually discrete, we will need to interpolate within the EDF of the reference channel to find the value of x'_1 , which must then be rounded to the nearest integer.

RESULTS AND DISCUSSION

A normalization look-up table was generated from a dependent sample consisting of unnormalized data from GOES-7 on 18 May 1988. The table was

Figure 4. Location on globe of dependent and independent samples of image data, 1996 × 2400 pixels.



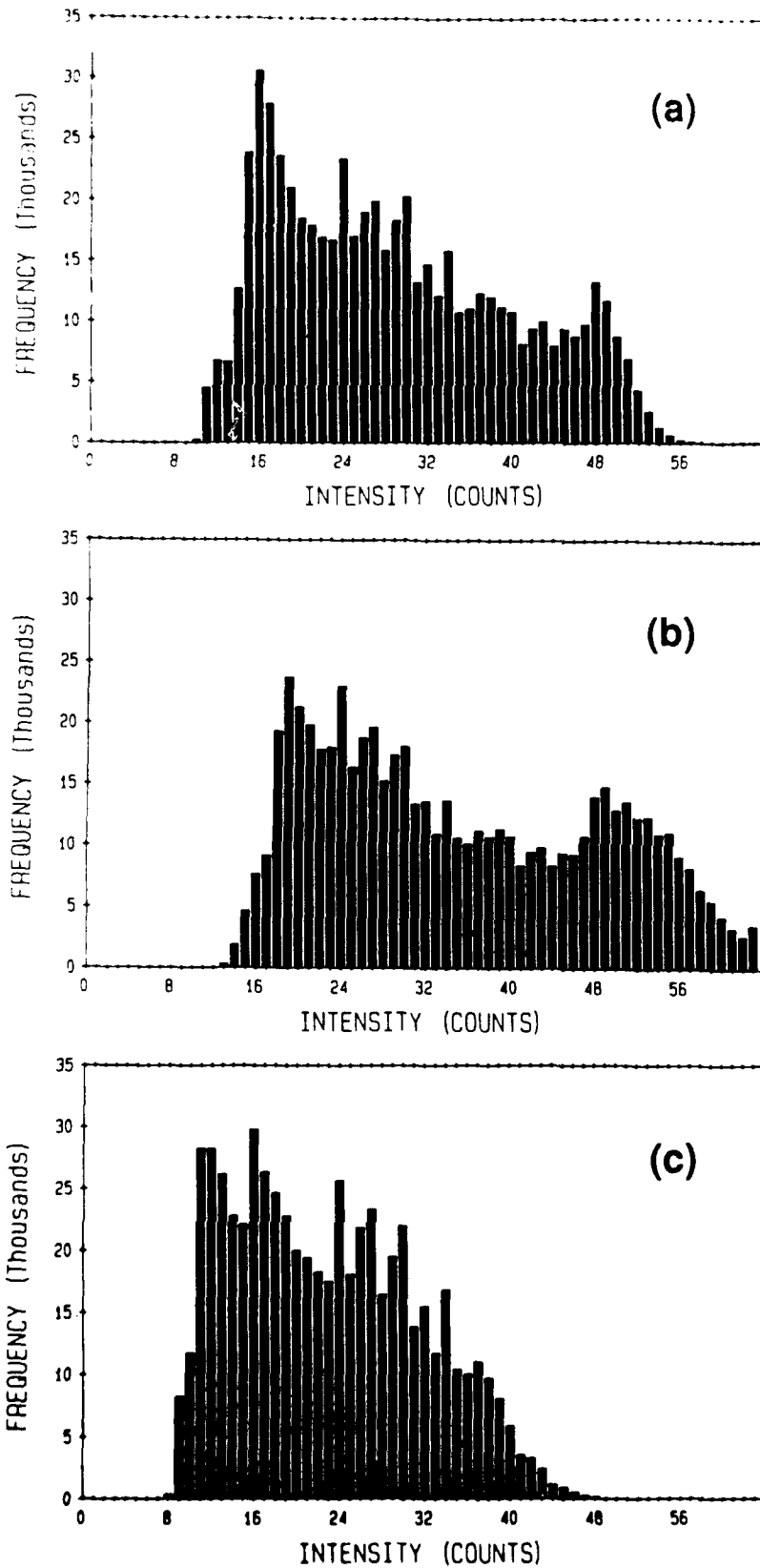


Figure 5. Histograms of unnormalized GOES-7 image data, 18 May 1988: (a) Channel 2; (b) Channel 5; (c) Channel 6.

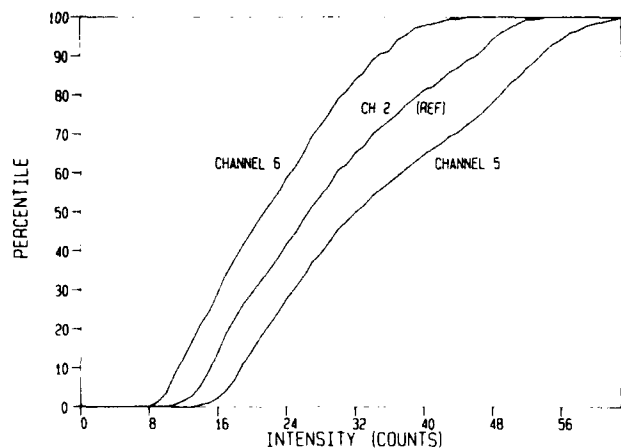


Figure 6. Empirical distribution functions for unnormalized GOES-7 image data, 18 May 1988.

applied to normalize an independent sample of image data from GOES-7 2 weeks later, on 1 June 1988. Both samples were produced at the same time of day, approximately 1400 EDT (1800Z), when the disk of the earth beneath the satellite was in full sunlight.

Generation of Normalization Look-Up Table

The dependent sample consists of unnormalized data from all pixels in a rectangular sector 1996 pixels east-west and 2400 pixels north-south, depicted at the equator in Fig. 4. An image of approximately one sixth of that sector appears in Fig. 2, and, as we saw previously, it is striped. The striping is most severe at the mid- and high-intensity levels (low and high cloud), which appear gray and white, respectively.

For each of the channels, histograms were compiled from the data in the full dependent sample. Each histogram represents data from 598,800 pixels. The three panels in Fig. 5 show the histograms for Channels 2, 5, and 6. (The abscissae, labelled "Intensity (Counts)," are the output levels.) Based on the criteria mentioned previously, Channel 2 was designated as the reference. We chose Channels 5 and 6 as examples because their histograms are the most different from the reference channel's. The histogram of Channel 5 is displaced towards the high intensities

Figure 7. Unnormalized GOES-7 image, 1 June 1988.



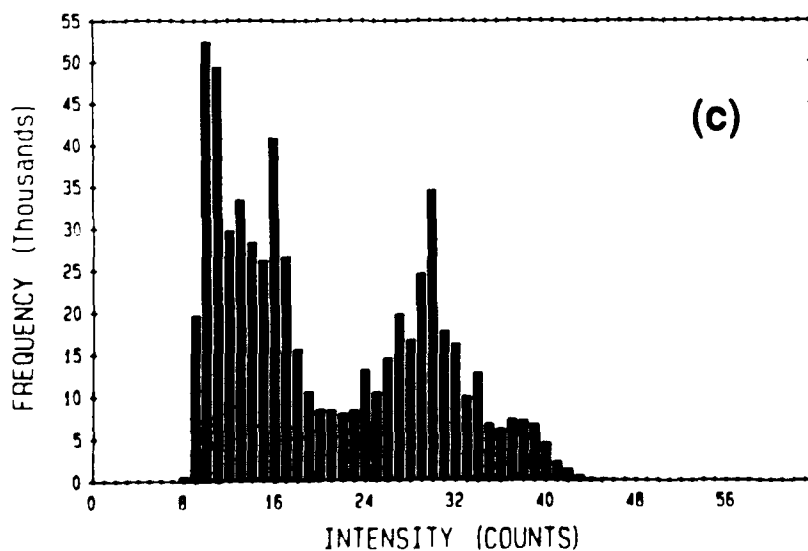
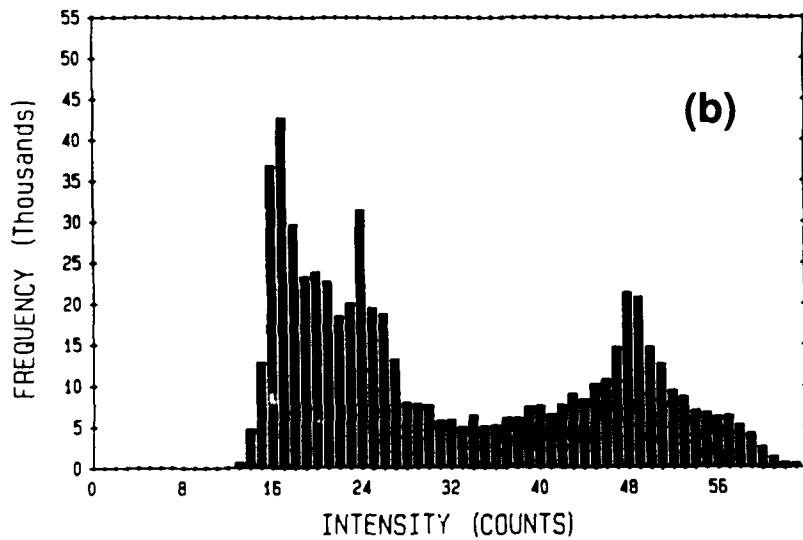
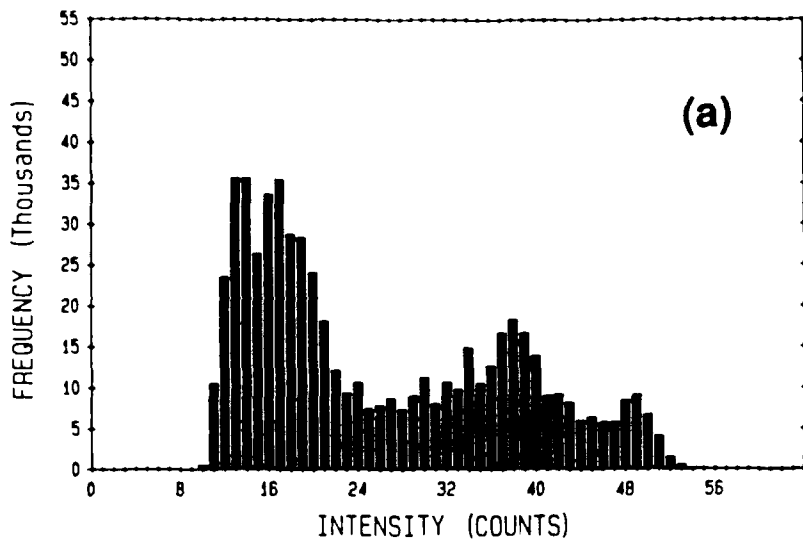


Figure 8. Histograms of unnormalized GOES-7 image data, 1 June 1988: (a) Channel 2; (b) Channel 5; (c) Channel 6.

relative to that of the reference channel, its upper end is clipped, and it is broader than the reference channel's, indicating that the gain in Channel 5 is greater overall than that of the reference channel. With Channel 6, the situation is reversed, because the gain is lower overall than that in the reference channel. The differences among the three histograms also suggest that the dependence of gain on intensity (or the degree of nonlinearity) varies from channel to channel.

Figure 6 shows the EDFs of the three channels, which were computed from the histograms in Fig. 5. The abscissa is the output level, and the ordinate is the percentage of the data with outputs at or below that level. We applied Eq. (2) to these EDFs and those of the other channels to generate the normalization look-up table. [See Weinreb et al. (1989) for a listing of the table.] That table will be applied to independent image data, as described below.

Application to Independent Sample

The independent sample of unnormalized GOES-7 data was produced on 1 June 1988, from the upper sector of 1996×2400 pixels depicted in Fig. 4. An image of approximately one sixth of that sector is shown in Fig. 7. As expected, there is striping, and it is most severe at the mid- and high-intensity levels. Figures 8 and 9 show the histograms and EDFs, respectively, compiled in Channels 2, 5, and 6. Their shapes are quite unlike the corre-

sponding histograms and EDFs from the dependent sector, reflecting the differences in the cloud distributions in the two sectors. However, for the normalization look-up table to be effective on the independent sector, the channel-to-channel relationships among the EDFs must be similar for the two sectors; i.e., the relative (channel to channel) gain functions must be similar. Although it may not be obvious at this point, the results below demonstrate that this is in fact the case.

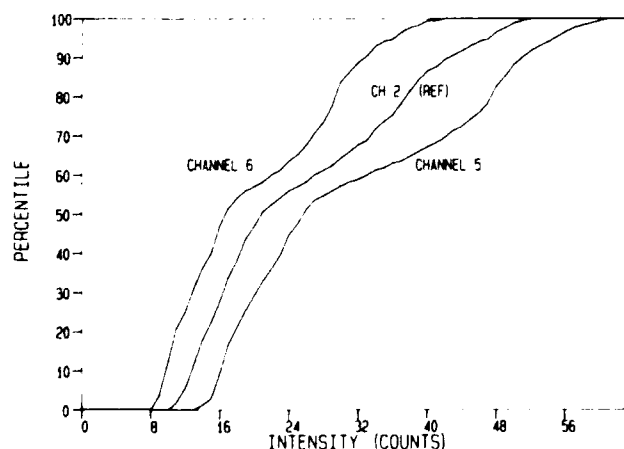
We normalized the 1 June data by applying the 18 May look-up table. Figure 10 shows the histograms of the normalized data for Channels 5 and 6. In position and shape they are now much more like the histogram of the reference channel (reproduced in the upper panel), as they should be if the normalization is to be successful. An unusual feature in these histograms is the presence of breaks, i.e., intensity levels with a zero frequency of occurrence. Because the data are discrete, breaks will occur when normalization expands a region of a histogram, as is explained more fully in Weinreb et al. (1989).

Figure 11 shows the EDFs of the normalized data for all three channels. As it should, the normalization process made the differences among them practically insignificant. (The EDFs of the other five channels behave similarly.) The largest differences occur where the EDFs of Channels 5 and 6 have flat spots. These are caused by the breaks in the histograms and are an artifact of digitization.

Figure 12 is the normalized image of the same area as was shown in Fig. 7. It is the "after" to the "before" of Fig. 7. The improvement is obvious, since we cannot see any stripes in the image.

The effectiveness of the normalization means that the channel-to-channel relationships among the EDFs remained essentially the same between 18 May and 1 June, as was surmised earlier, even though the cloud patterns and the EDFs themselves changed. Therefore, the relative gains and offsets among the channels also must have remained the same. Furthermore, although we do not show the results here, we found that the effectiveness of the 18 May normalization look-up tables decreased only slightly with time over a 6-week period. Therefore, any changes in relative gains and offsets, which, e.g., might have resulted from seasonal variation of temperatures on the satellite, had to be small over that period.

Figure 9. Empirical distribution functions for unnormalized GOES-7 image data, 1 June 1988.



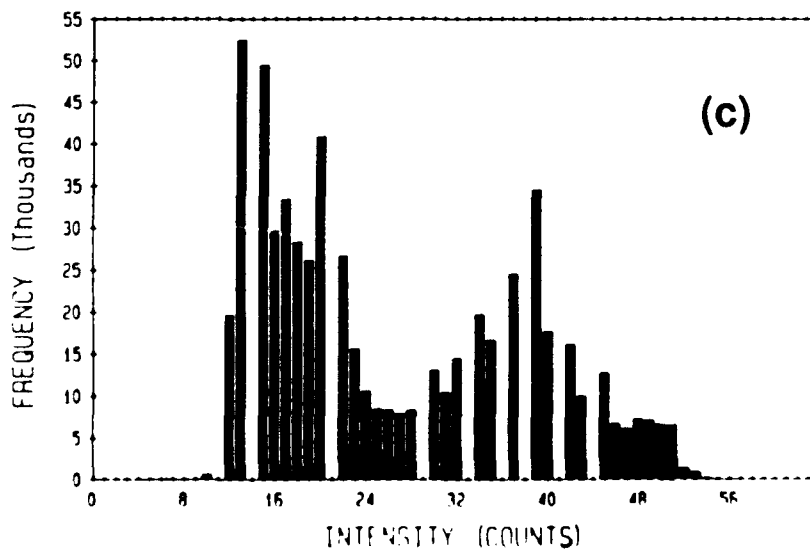
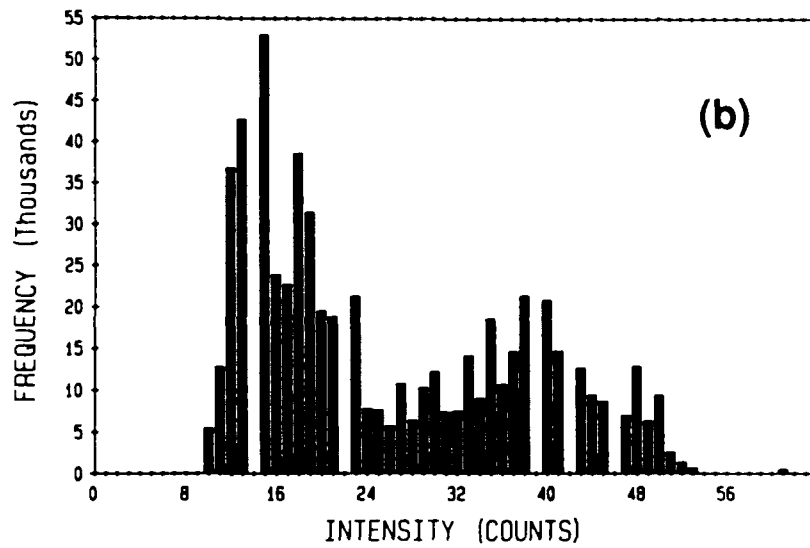
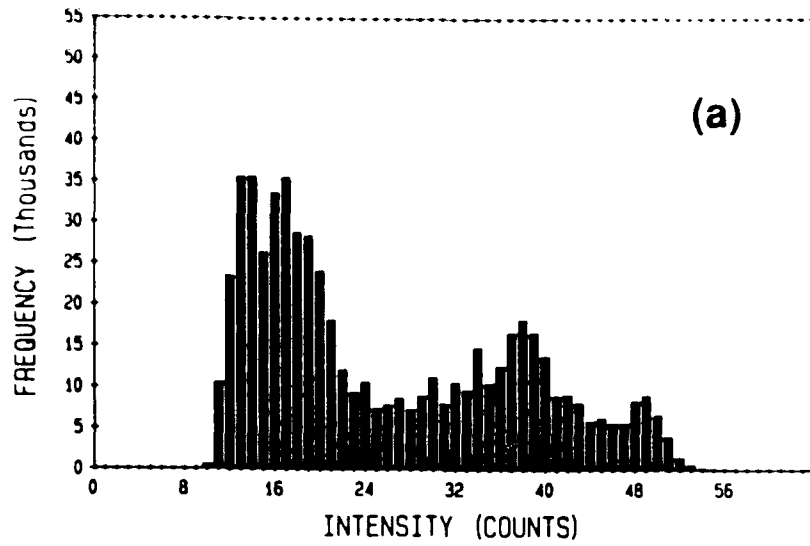


Figure 10. Histograms of normalized GOES-7 image data, 1 June 1988: (a) Channel 2; (b) Channel 5; (c) Channel 6. (Data in Channel 2, the reference channel, are never normalized.)

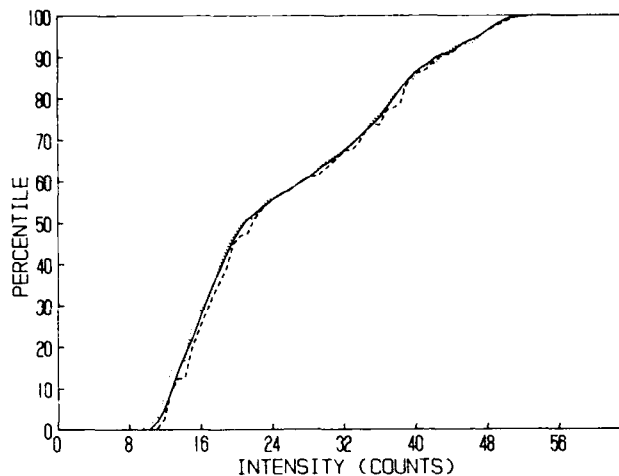


Figure 11. Empirical distribution functions for normalized GOES-7 data, 1 June 1988: (—) Channel 2 (reference); (···) Channel 5; (---) Channel 6.

CONCLUSION

The case study presented in this paper is strong evidence that normalization by EDF matching is an effective method for removing striping from visible images from GOES. Application of a normalization look-up table generated from GOES-7 data of 18 May 1988 removed the stripes from an image obtained on 1 June 1988. The method worked despite the nonlinearities in the outputs of the GOES-7 visible channels and the substantial channel-to-channel nonuniformities in gain. We expect it to work at least as well with data from GOES I-M, because the responses in the GOES I-M channels are expected to be linear and more nearly uniform than those of GOES-7. Further, we found that the same normalization table remained effective for at least 6 weeks. Since the gains of the GOES I-M channels are expected to be more stable in time than those of GOES-7, we would

Figure 12. Normalized GOES-7 image, 1 June 1988.



expect a normalization table to remain valid for even longer periods with GOES I-M.

The authors are indebted to Dong Han of NOAA/NESDIS for providing the raw image data; to Gene Legg of NOAA/NESDIS for producing the "sectorizer" images that appear in Figs. 2, 7, and 12; to Donald Mack of Integral Systems, Inc., of Lanham, MD, for supplying information on the application of EDF matching to data from Landsat; and to George Jones and Paige Bridges of NOAA/NESDIS for preparing the figures.

REFERENCES

- Bristor, C. L. (Ed.) (1975), Central processing and analysis of geostationary satellite data, NOAA Tech. Mem. NESS 64, National Oceanic and Atmospheric Administration, National Environmental Satellite Service, Washington, DC, 155 pp.
- Ensor, G. J. (1978), *User's Guide to the Operation of the NOAA Geostationary Satellite System*, U. S. Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite Service, Washington, DC, October 1978, 101 pp.
- Horn, B. K. P., and Woodham, R. J. (1979), Destriping LANDSAT MSS images by histogram modification. *Comput. Graphics Image Process.* 10:69-83.
- Komajda, R. J., and McKenzie, K. (1987), An introduction to the GOES I-M imager and sounder instruments and the GVAR retransmission format, NOAA Tech. Rep. NESDIS 33, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, Washington, DC, 51 pp.
- Weinreb, M. P., Xie, R., Lienesch, J. H., and Crosby, D. S. (1989), Removing stripes in GOES images by matching empirical distribution functions, NOAA Tech. Mem. NESDIS 26, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, Washington, DC, 25 pp.

1 For	
GOES I	<input checked="" type="checkbox"/>
GOES II	<input type="checkbox"/>
GOES III	<input type="checkbox"/>
Distribution/	
Availability Codes	
Avail and/or	
Dist	Special
A-1	20